



Austin Ward

ADAMSON AHDOOT, LLP

Defending client-AI communications in California discovery

A PLAINTIFF'S GUIDE TO THE ATTORNEY-CLIENT PRIVILEGE AFTER THE FEDERAL COURT RULING IN *HEPPNER*, A CASE OF FIRST IMPRESSION

Your client has already used AI. Before she reached your conference room, she typed the facts of a car crash into ChatGPT, asked Claude to organize a sprawling treatment chronology, or used an AI assistant to turn confusing medical records into a list of questions for counsel. She brought that output to your first meeting. You used it. It shaped your strategy.

Defense counsel will soon serve a request for production seeking “all communications with any artificial intelligence platform related to your claims in this case.” They will cite *United States v. Heppner* (S.D.N.Y., Feb. 17, 2026, No. 25 Cr. 503 (JSR)) ECF No. 27, a first-of-its-kind federal ruling holding that a criminal defendant’s exchanges with Anthropic’s Claude were protected by neither attorney-client privilege nor the work product doctrine. Judge Rakoff described the question as one of “first impression nationwide.”

That opinion is serious and will be cited across the country. But it is a federal criminal decision applying federal common law, on facts unusually bad for the privilege claim. California law offers a stronger basis for resisting waiver arguments. While no California appellate court has yet squarely resolved whether consumer large language model (LLM) inputs fall within privileged protections, California starts from a different place than federal law. Our privilege law is statutory. Its text is broader, more concrete, and more protective than the framework Judge Rakoff applied.

As explained below, California lawyers should resist the temptation to argue that every client interaction with a generative AI platform is, or is not, privileged. Not every use of AI warrants the same treatment. The strongest claims of privilege will rest where AI is used to organize facts, summarize records, draft questions, or otherwise facilitate commu-

nication with counsel – especially when counsel directs that use. On the other hand, the weaker and more vulnerable cases are those in which a client independently uses a consumer AI platform as a substitute for legal advice, while having retained separate counsel.

This article explains how *Heppner* will be used in discovery, why it does not provide the appropriate framing for this issue, why California law provides stronger answers than federal common law, and where the best limiting principles lie for plaintiff-side lawyers defending claims of privilege involving AI-related communications.

What *Heppner* held

Bradley Heppner, the former chair of a publicly traded company, was indicted on securities fraud and related charges stemming from an alleged \$150 million investor fraud. Without any direction from counsel, he used the consumer version of Claude to prepare reports about possible defense arguments and strategy. He later shared those materials with his lawyers at Quinn Emanuel. Counsel acknowledged the AI-generated documents “affect[ed]” strategy but conceded they did not “reflect” counsel’s strategy at the time Heppner created them.

Judge Rakoff rejected privilege on three grounds: (1) Claude is not an attorney, and privilege requires a “trusting human relationship” with a licensed professional; (2) Anthropic’s privacy policy, which permits data collection, model training, and disclosure to government authorities, destroyed any reasonable expectation of confidentiality; and (3) Heppner did not use Claude for the purpose of obtaining legal advice from his attorney. The work product claim failed because the materials were not prepared by or at the direction of counsel.

On the same day, a federal court in Michigan reached the opposite conclusion on work product. In *Warner v. Gil-*

barco, Inc. (E.D. Mich., Feb. 10, 2026, No. 2:24-cv-12333), Magistrate Judge Patti held that a pro se litigant’s ChatGPT-assisted materials were protected opinion work product, reasoning that generative AI programs “are tools, not persons” and that inputting materials into an AI platform did not waive protection. The *Heppner/Warner* split frames the central question: Are inputs to an AI platform disclosures to a third party or mere utilization of a tool?

AI is not an attorney; it is a tool

Privilege requires a communication between client and attorney, and Claude is not an attorney. Judge Rakoff thus held: no privilege can attach to a communication with a non-lawyer.

Evidence Code section 954 protects “a confidential communication between client and lawyer.” (All subsequent statutory references are to the California Evidence Code.) “The privilege embraces not only oral or written statements but actions, signs, or other means of communicating information by a client to his attorney.” (*City & County of San Francisco v. Superior Court* (1951) 37 Cal.2d 227, 235, emphasis supplied.) It is well-settled that confidential writings (e.g., notes, memoranda) a client creates for later transmittal to an attorney in connection with anticipated litigation remain privileged.

While the analysis performed by Magistrate Judge in *Warner* concerned the work product doctrine as applied to a pro se litigant, it is relevant he found “ChatGPT (and other generative AI programs) are tools, not persons, even if they may have administrators somewhere in the background.” It is this mode of reasoning that may characterize LLMs as sophisticated writing aids within the user’s drafting environment.

Electronic writing assistance has taken many forms over the past decade, increasingly providing more active and sophisticated linguistic editorial supports

beyond spell-check technologies. On mobile devices, for example, predictive text engines learn from the user's writing patterns, including the content of prior messages and larger datasets drawn from other users, to predict and generate entire words and phrases before the user finishes typing. Grammar and style assistants like Grammarly go further still. Grammarly ingests the entirety of whatever the user is writing – emails, letters, legal memoranda – transmits it to cloud servers, runs it through machine learning models trained on user data, and returns suggestions for rewording, restructuring, and tone. The platform stores user content on remote servers, and its privacy policy reserves the right to use anonymized and aggregated data for product improvement.

Generative AI tools build on these systems and offer clients increased capacity for assimilating and organizing large quantities of data for later transmission to counsel. This can look like systematizing medical records, billing records, communications, photographs, tax documentation, and other related materials. In addition to organizing the materials themselves, the tools empower the user to search the content of documents and things and arrange them in ways that facilitate effective communication with counsel (e.g., identifying relevant datasets the client wishes to isolate and convey). Clients, in an effort to aid in communicating with their attorney, may prompt the technology to create medical chronologies, parse medical records for treatment obtained prior to and after an incident, identify which medical bills have been paid and which remain outstanding, pull out the names, addresses, and phone numbers of witnesses listed in incident reports and police reports, among varieties of similar clerical tasks.

ChatGPT and Claude accomplish this by processing user input through language models and return substantive responses – summaries, reorganized content, answers to questions – rather than corrected spelling or predicted phrases. But the underlying mechanic is the same: the user's text is transmitted to a remote server, processed by a machine learning

model, and used to generate a response. The data handling – cloud transmission, server-side processing, model training on aggregated data, retention under terms of service – is structurally analogous to what Grammarly and related services do.

However, the output programming of generative AI tools creates novel information environments beyond *mere* improvement on what the user has written or intends to communicate. These AI conversations may affirmatively *create* output that resembles advice, new information, and instruction generated by third-party algorithms and massive independent datasets. Moreover, the level of sophistication available for this novel information and output gestures forcefully toward the presence of a third-party dynamically interacting with the client to an extent readily distinguishable from Grammarly or advanced predictive-text capabilities.

Where the content of output appears to provide novel advice and information to the user – absent involvement from counsel – it is less likely to be treated as privileged. The activity becomes too similar to the act of performing substantive independent research (e.g., Google), which will ordinarily be considered unprivileged.

Those seeking disclosure of client-AI communications will argue that the involvement of third parties in this context exceeds acting as a passive conduit of the communication to counsel and presents as a far more active participant in the communication. This may be true. However, there is no requirement that preserving the privilege requires limiting disclosures to mere “passive conduits” of the communication. California explicitly permits preserving the privilege of attorney-client communications where disclosure is mediated by certain types of third parties actively engaged in the communication; California courts already apply this logic to protect communications involving interpreters and family members who assist vulnerable clients. (*De Los Santos v. Superior Court* (1980) 27 Cal.3d 677, 683–685.)

Section 952 anchors the definition of attorney-client communications to include communications made through persons “to whom disclosure is *reason-*

ably necessary for the transmission of the information *or the accomplishment of the purpose* for which the lawyer is consulted.” (Emphasis supplied.) Section 912, subdivision (d) repeats the same language: a disclosure of communications in confidence that is “reasonably necessary for the accomplishment of the purpose for which the lawyer ... was consulted” does not waive the privilege. Thus, whether a client's conversations with generative AI tools can be deemed an attorney-client communication will turn on the extent the communication was “reasonably necessary” to “accomplish[] the purpose for which the lawyer is consulted.” Even where the communication serves a dual purpose, the privilege applies if the predominating purpose is transmission to counsel. (*Holm v. Superior Court* (1954) 42 Cal.2d 500, 507.)

When a plaintiff uses an AI tool to organize scattered medical facts or frame questions about her legal rights, the platform might be understood to serve a reasonably necessary facilitative function within sections 952 and 912's safe harbor.

There are policy considerations that favor adopting a more liberal application of the predominant purpose standard when employed by an unsophisticated client. Ordinary consumers enter every legal dispute at a structural disadvantage compared to business clients. At the very first meeting with counsel, many people struggle to articulate the relevant facts, identify which of their experiences sound in recognized causes of action, organize stacks of hard-copy and electronic documents, or convey their circumstances with the coherence that effective representation demands. Corporate defendants and institutional insurers face none of these barriers. They understand basic legal frameworks, maintain in-house platforms that electronically organize vast quantities of data for immediate recall, and arrive at counsel's office prepared to cooperate efficiently. LLM technologies have the potential to close that gap, leveling the floor of access to justice – to give an overwhelmed individual the same ability to organize, synthesize, and communicate that a corporate client has.

The threshold question under Cali-

California law is not whether the AI platform is a licensed attorney. It is whether the predominant purpose of the client's writings on the platform was for confidential communications to counsel, and whether utilizing the platform was reasonably necessary for the accomplishment of the purpose for which the lawyer is consulted.

Disclosure to third-party electronic vendors with "access" to the content of writings should not per se waive privilege

Heppner relied heavily on Anthropic's privacy policy, which permits collection of user inputs, retention for model training, and disclosure to third parties including government authorities. Defense counsel will invoke that reasoning to argue that a client who types case-related facts into an AI platform has consented to disclosure and therefore surrendered confidentiality. California law supplies a more textually grounded response, but the analysis should proceed in distinct steps.

Evidence Code section 917(b)

California's legislature has addressed this concern directly. Evidence Code section 917, subdivision (b) provides that a communication between privileged persons "does not lose its privileged character for the sole reason that it is communicated by electronic means or because persons involved in the delivery, facilitation, or storage of electronic communication may have access to the content of the communication." This statute does three things. First, it establishes that electronic transmission does not, standing alone, defeat privilege. Second, it addresses the concern *Heppner* relied on – vendor access – by providing that privilege is not defeated because persons involved in the delivery, facilitation, or storage of electronic communication *may have access* to the content. Third, it works with the presumption of confidentiality in section 917(a), which places the burden on the party opposing privilege to prove the communication was not confidential if it was made in confidence "*in the course of the lawyer-client . . . relationship.*"

Section 952's subjective standard

Even if third-party access could work to waive privilege, California does not apply a "reasonable expectation of priva-

cy" test for attorney-client privilege. That concept belongs to Fourth Amendment jurisprudence, not the Evidence Code. Section 952 asks whether the communication was made "by a means which, *so far as the client is aware*, discloses the information to no third persons." (Emphasis supplied.) That phrase establishes a subjective standard focused on the client's actual awareness, not on what a hypothetical reasonable person would understand after reading a terms-of-service document. Peer-reviewed research consistently demonstrates that fewer than one percent of users read TOS documents. A click-through agreement the client never read is weak evidence of actual subjective awareness.

The defense will respond with constructive notice. That is true as contract law. But section 952 does not say "so far as the client should have been aware" or "so far as a reasonable person would have been aware." It says "is aware."

If the argument advanced by the defense is not merely that a third-party vendor had access, but too much access – or, not merely that the third-party vendor used the data but distributed the information to other entities – then strict application of a subjective standard for an unsophisticated client, undoubtedly ignorant of these nuances, becomes all the more critical to preserve.

The technology reduction

If *Heppner's* reasoning were accepted, the consequences would destabilize the entire infrastructure of modern legal practice. Every privileged email sent through Gmail or Outlook was never confidential, because the provider's terms authorize data processing and legal-process compliance. Every document stored in cloud platforms was never protected. Every privileged conference conducted over Zoom or Teams was compromised.

These platforms all have terms of service that reserve rights of access and use. Many have mandated reporting requirements for certain illegal conduct requiring active monitoring of content. Google, for example, uses hash-matching and machine-learning classifiers to detect illegal material in Gmail, analyzing behavioral patterns (i.e., suspicious mes-

saging and interaction trends), and has produced emails under search warrants and subpoenas on countless occasions. Microsoft's consumer-tier Copilot terms explicitly warn users not to share information they do not want reviewed. Every major smartphone voice assistant has been documented having human contractors listen to recordings of private conversations. Terms of service often explicitly reserve the right to do this.

No California court has endorsed the conclusion that third-party vendors' mere access to electronic communications waives privilege. The legal profession operates on the assumption that technological intermediaries are tools, not third parties, for purposes of privilege and that assumption has been validated by decades of ethics opinions, case law, and statutory provisions including section 917(b). The California State Bar's Formal Opinion 2010-179 confirmed that cloud storage does not per se destroy privilege, and Formal Opinion 2012-184 extended that analysis to virtual law offices.

While attorneys have heightened duties to protect the security of confidential communications by performing due diligence that third-party vendors having access to confidential work product and communications demonstrate sufficient competence and interest in maintaining the confidentiality of those materials, clients do not bear the same duties. Even where communications are conducted over electronic platforms that do not adequately secure the confidentiality of the communication, section 952's subjective standard should still work to protect a client who earnestly believed the communications would remain adequately secured.

The training distinction

The most prominent limiting principle for AI platforms tends to focus on model training: When an AI platform uses inputs to train its models, the communication is arguably incorporated into the model's parameters and may influence outputs to other users.

Model training derives statistical patterns from massive datasets and is not designed to make individual prompts searchable. Repeatedly, AI companies represent that no user can query the

training model to retrieve a specific client's inputs. Expert technical assessments of precisely what happens with the information in the context of training models may be necessary to more cogently articulate the rational basis for this as the predicate for warranting waiver. If these companies are to be believed (granted, a large *if*), then the use of inputs into training models would be tantamount to shredding a document and mixing the fragments with millions of other shredded documents and saying the material was "processed" into the model. No one can reconstruct the original, and the confidentiality of the original content is preserved.

This industry narrative has important gaps. LLMs sometimes memorize and reproduce fragments of training data, particularly when inputs are unusual or repeated. This is a known failure mode called "memorization" that occurs at low rates, that AI companies actively work to mitigate, and that produces fragments rather than coherent communications. Importantly, this is leakage, not intentional disclosure. California privilege law has never treated the theoretical possibility of inadvertent disclosure as equivalent to actual disclosure – e.g., restaurant conversations, phone calls over a cellular network, security vulnerabilities in a cloud service. The standard isn't zero risk of exposure; it's whether the client made the communication by means that, so far as the client was aware, disclosed it to no third person.

Regardless, section 917(b) states the privilege is not lost because "persons involved in the . . . storage of electronic communication may have access to the content." The word "access" is doing work here. In one sense, training is *less* access than what 917(b) already permits. The vendor's engineers can't pull up a specific user's session from the trained model the way a cloud storage provider can pull up a specific file.

The *training* distinction is increasingly untenable as a principled line while Google uses Gmail data to improve spam filters, Microsoft uses enterprise data for service improvement, and Apple processes Siri commands for quality improve-

ment – none of which has been held to defeat privilege.

Counsel-directed AI use

The preceding sections address the common scenario: A client who independently uses an AI platform and later shares the output with counsel. But the strongest privilege and work product position arises when counsel *directs* the client to use the platform in the first place.

Attorney-client privilege

When counsel instructs a client to use an AI service, that instruction reframes the entire interaction as one undertaken to accomplish the purpose of the legal consultation. California's "reasonably necessary" standard under sections 952 and 912, subdivision (d) is notably more protective than the federal "nearly indispensable" test. (*Conway v. Licata* (D. Mass. 2015) 104 F.Supp.3d 104.) When an attorney instructs a client to use AI to organize scattered medical records or prepare a factual narrative for counsel's review, the platform's involvement serves the consultation's purpose just as directly as a translator or paralegal would. This answers the purposive weakness that sank the privilege claim in *Heppner*: the client's AI use becomes an extension of the attorney-client relationship from the moment the instruction is given.

Work product doctrine

Under Code of Civil Procedure section 2018.030, subdivision (a), writings that reflect an attorney's impressions, conclusions, opinions, or legal theories are absolutely protected. (*Southern California Edison Co. v. Superior Court* (2024) 102 Cal.App.5th 573.) When carried out at the direction of counsel, the input into the AI platform would directly reflect the attorney's impressions, conclusions, opinions or legal theories, and be entitled to absolute protection.

Work product "is not waived except by a disclosure wholly inconsistent with the purposes of the privilege, which is to safeguard the attorney's work product and trial preparation." (*OXY Resources California LLC v. Superior Court* (2004) 115 Cal.App.4th 874, 891.) California courts have established that work product waiver requires a higher threshold than attorney-client privilege waiver because

the work product doctrine protects against "opposing parties, rather than against all others outside a particular confidential relationship." (*BP Alaska Exploration, Inc. v. Superior Court* (1988) 199 Cal.App.3d 1240.)

The line becomes whether disclosure was made to a person who has "no interest in maintaining the confidentiality . . . of a significant part of the work product." (*OXY Resources California LLC, supra*, 115 Cal.App.4th at 891.) Thus, the "interest in maintaining the confidentiality" of the work product becomes the crux.

Mere disclosure of work product to third parties does not waive work-product protection. If it did, the doctrine would foreclose the use by courts and practitioners of some of the most routinely relied upon technologies for legal work in the United States. Every legal research query submitted to Westlaw, Lexis+, or CoCounsel would constitute a disclosure of work product. Thomson Reuters' Westlaw states in its privacy policy that it collects, uses, and, in some circumstances, shares information provided to it. Enterprise agreements may include additional terms that promise confidentiality as to work product, but no service guarantees that the provider will have no access whatsoever to the underlying content.

For now, opposing parties' mutual dependence on third-party vendors for hosting, storing, and processing attorney work product has created something of an armistice, with minimal litigation pressing seriously into the matter. If that truce ends, *OXY Resources California LLC* will require trial courts to determine on a case-by-case basis whether a third-party vendor's interest in maintaining the confidentiality of disclosed work product is sufficiently strong to defeat waiver. That inquiry will be fact-intensive: Parties seeking to preserve protection will need to marshal evidence of the vendor's security features, contractual confidentiality obligations, corporate competence, and constraints on the use and distribution of content.

The Northern District of California has already recognized work-product protection for AI-related materials prepared by counsel. In *Tremblay v. OpenAI, Inc.* (N.D. Cal. 2024) 2024 WL 3748003, the

court held that unused prompts, account data, and testing results constituted opinion work product. While not discussing client communications with AI products, the rationale used in *Tremblay* should apply with equal force in circumstances where the client is acting at the behest of counsel.

While utilizing AI tools managed by vendors with a legitimate interest in maintaining confidentiality should not per se result in waiver, the ethical duties attorneys have in managing confidential information, even where it does not amount to waiver, remains properly ripe for further discussion and subject to diligence concerning the actual quality of security afforded by third-party vendors amidst these burgeoning technologies.

The stakes

Attorney-client privilege exists to encourage full and frank communication between clients and lawyers. (*Costco Wholesale Corp. v. Superior Court* (2009) 47 Cal.4th 725, 732.) For many plaintiff clients, AI is not a substitute for counsel – it is a bridge to counsel. It helps them assemble dates, summarize treatment, frame questions, and describe what

happened coherently before they sit down with a lawyer. A rule that threatens discovery of that preparation does not protect the attorney-client relationship. It chills it.

It also creates an unjustifiable class divide. The sophisticated client who arrives with a neat handwritten chronology is safe. The overwhelmed client who needs technological help to accomplish the exact same work is punished for using the available tool. California courts should be deeply reluctant to endorse that result.

When the defense serves that request for production, do not concede the premise. Separate privilege from work product. Separate raw chats from communications transmitted to counsel. Separate self-directed consumer use from attorney-directed preparation. Force the court to engage with California's statutory text. And if the court intends to find waiver by disclosure to third parties with access to the content, press on the need for a meaningful limiting principle that distinguishes the monitored content in Claude from the monitored content in Google Docs.

Heppner matters because it is first. But if we want California law to develop differently, we need to make the argument now, on a clean record, before *Heppner* hardens into conventional wisdom. 

Austin Ward is a partner at Adamson Ahdoot, LLP practicing in tort litigation with an emphasis in personal injury. He graduated from Pepperdine University School of Law (Class of '16). Before that, he attended Westmont College where he earned two bachelor's degrees in Philosophy and Reasoning & Advocacy.